

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
11 July 2002 (11.07.2002)

PCT

(10) International Publication Number  
**WO 02/054258 A1**(51) International Patent Classification<sup>7</sup>: **G06F 13/00**

(21) International Application Number: PCT/US02/00263

(22) International Filing Date: 4 January 2002 (04.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/259,651 4 January 2001 (04.01.2001) US(71) Applicant (for all designated States except US): **RAG-NAROK SYSTEMS, INC.** [US/US]; 230 South LaSalle Street, Suite 688, Chicago, IL 60604 (US).

(72) Inventor; and

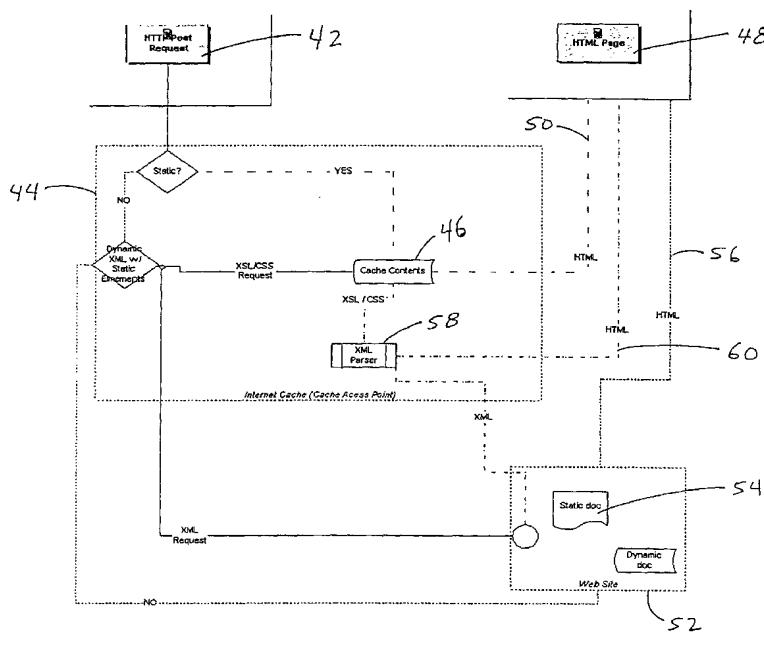
(75) Inventor/Applicant (for US only): **PAWEJKO, Tomasz** [US/US]; 3132 N. 77th Court, Elmwood Park, IL 60707 (US).(74) Agent: **MENCHACA, Alejandro**; McAndrews Held & Malloy, Ltd., 500 W. Madison, Suite 3400, Chicago, IL 60661 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

*[Continued on next page]*

(54) Title: METHOD AND SYSTEM FOR CACHING AND DISTRIBUTING INFORMATION FROM DYNAMIC INTERNET WEBSITES



(57) Abstract: A method and system for universally caching and distributing information (48) dynamically-generated websites (48) is provided. The method comprises transmitting a request (42) for a webpage and processing (44) the request such that the static (54) and dynamic components of the requested webpage are parsed and retrieved separately and then combined for distribution to the location from which the request was first transmitted. The system comprises an input device for receiving the request for the webpage and a cache access point for processing the request by parsing (58) and retrieving separately the static (54) and dynamic components of the webpage and then combining those components for distribution to the location (42) from which the request was first transmitted.

**WO 02/054258 A1**



**Published:**

- *with international search report*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**METHOD AND SYSTEM FOR CACHING AND DISTRIBUTING  
INFORMATION FROM DYNAMIC INTERNET WEBSITES**

**BACKGROUND OF THE INVENTION**

The invention relates to dynamically-generated internet web sites and a method for universally caching and distributing information from those dynamically-generated web sites.

In utilizing the world wide web, user's often seek access to numerous types of web pages. Two standard types of web pages are those that are static with very little of their content changing over time. Other web pages are dynamically-generated, that is, the information displayed on those web sites constantly changes from one moment to the next.

One major impediment to satisfactory use of the world wide web is the time delay between the request for a web page and the display of that web page on the user's computer screen. In connection with static web pages, one method employed to reduce the time between request and display is to cache the web site locally in the computer.

A second method to reduce the time for display is to retrieve the static web page not directly from the web site, but from the internet service provider or other local internet nodes where the static web page may be cached. Substantial caching is often accomplished by the use services obtained from Mirror Image, Inc.

Because dynamically-generated web sites are

constantly changing, any cached pages will become obsolete almost immediately. Thus, when requested by a user, the web page cannot be retrieved with the most current information from any cached site. To retrieve the most current information, the user's request must obtain the most current web page directly from the web site. This increases general internet traffic and traffic to the web site. In turn, requiring direct access to the web site delays the display of the web page on the user's computer screen.

One method of addressing this problem is separation of data from formatting and processing instructions. This functionality is built into the program offered by Microsoft under the name Internet Explorer. Microsoft's Internet Explorer includes the MS XML Parser. The MS XML Parser enables Internet Explorer to process data in the form of XML documents; which in turn enables utilization of cached static portions of a dynamically-generated web page in combination with the dynamic data to complete the display of the dynamically-generated web page. That is, the MS XML Parser can access cached web pages, parse them and apply those static portions of the web page in displaying a dynamic web page. Only the dynamic data must be retrieved directly from the web site.

Examples of static information on dynamically-generated web pages include formatting schemes, scripts, processing instructions, banners, font information, and

the like. Examples of dynamic information include changing data specific to the web page, such as stock quotes for a financial web page.

The use of the MS XML Parser, however, is only available for use with Microsoft's Internet Explorer 5.0 on a user's computer. Although the use of Microsoft's Internet Explorer 5.0 and the MS XML Parser reduce the amount of time necessary to display a dynamically-generated web page, there exist undesirable limitations of this system. The first limitation is that the MS XML Parser is not available for use with other web browser software such as earlier versions of Internet Explorer and Netscape Navigator 5.0 and earlier versions. A second limitation is that even if one uses Microsoft's Internet Explorer 5.0, the limited amount of memory available on the user's computer for caching web page information provides only a limited caching capacity much smaller than that present at the internet service provider of the various internet nodes where caching is accomplished. Thus, because of the limited amount of cache memory for web pages on a user's computer, the MS XML Parser quickly loses its utility even with Microsoft Internet Explorer 5.0.

**SUMMARY OF THE INVENTION**

Accordingly, it is an object of the present invention to reduce the time between request by a user of a dynamically-generated web page and the display of that web page on the user's computer screen.

It is another object of the present invention to make the benefits of the MS XML Parser universally available to any web browser software and thereby reduce the time between dynamically-generated web page request by a user and display of the web page on the user's computer screen.

It is another object of the invention to make the MS XML Parser available in a location other than the user's computer.

It is another object of the invention to coordinate the use of the MS XML Parser with the caching capabilities of internet service providers and other internet nodes.

These and other objects of the invention are achieved in a method and system that utilizes the caching capabilities of internet service providers and other internet nodes in connection with dynamically-generated web pages. The system includes utilizing the standard caching systems presently available for static web pages.

The system further includes adapting and applying the MS XML Parser to the cached web page information located in the memory of the internet service provider or other internet node.

The method includes directing a user's request to a

getter. The getter then accesses the cache of the internet service provider or local internet node to determine whether the cache contains a copy of the requested dynamically-generated web page. If the cache does not contain a copy of the dynamically-generated web page, the getter determines that all of the contents of the web page must be obtained directly from the web site.

If the getter determines that the cache does contain a copy of the dynamically-generated web page, the MS XML Parser parses the cached web page to make available those components of the web page that are static. The static information is then combined with the dynamic information obtained directly from the web site to compile a complete dynamically-generated web page for display on the user's computer screen.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

For a more complete understanding of this invention reference should now be had to the embodiment illustrated in greater detail in the accompanying drawings and described below by way of example of the invention.

In the drawings:

FIG. 1 is block diagram of a system of the prior art.

FIG. 2 is block diagram of a system of the present invention.

**DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT**

Figure 1 is a block diagram of a system 10 of the prior art. The internet is typically accessed through a computer. The computer can be used to make requests to view particular internet pages and then to display those pages. In Figure 1, the request 12 is posted at a computer. The request 12 is transmitted through the user's internet connection (e.g., cable or modem connection) to the internet. The internet consists of an enormous network of databases and connections between those databases.

Additionally, the internet includes cache access points 16. The user's request 12 will typically encounter such a cache access point 16 prior to actually connecting to the web site 18 containing the web page sought in the request 12.

At the cache access point 16 the request 12 is evaluated to determine whether it is actually requesting a static web page or a web page generated from dynamic information. If the request 12 requests a web page generated from dynamic information and data, the request 12 must complete its connection to the web site so that the requested web page can ultimately be displayed at the user's computer 14.

If, however, the request 12 requests a static web page, the cache access point 16 may have that static web page cached in its data banks. Accordingly, the cache

access point 16 evaluates whether the web page of request 12 is cached in its cache contents 20. If the web page of request 12 is not cached in the cache contents 20, the request must connect to the web site 18 to obtain the static document 24. The static document 24 is typically then routed through the cache access point 16 so that the static document can then be cached in the cache contents 20 of the cache access point 16. The static document can then be routed to the user's computer 14 for display on the monitor of the user's computer 14.

If the request 12 is for a web page with dynamic data, the dynamic document 26 obtained from web site 18 is not routed through the cache access point 16 and is instead directly routed to the user's computer 14 for display on the monitor of the computer 14. In the system 10 of the prior art, the cache access point 16 will only recognize a web page of request 12 as static if all of the components of the web page of the request 12 are static. That is, if any component of the web page of request 12 requires dynamic data, then the cache access point 16 recognizes that web page as not static and will not access the cache contents 20.

Figure 2 is a block diagram of a system 40 of the present invention. The present invention recognizes that a request 42 is made by a user at a computer. The request 42 is routed through a cache access point 44. The cache access point 44 determines whether the web page of request

42 is completely composed of static data.

If the web page of request 42 is completely composed of static data, then the request 12 is processed in a manner identical to that of the prior art system 10 of Figure 1. That is, the cache access point 44 determines whether the web page of request 42 is cached in its cache contents 46. If the web page of request 42 is cached in the cache contents 46, then that web page is directly routed to the user's computer 48 for display on the monitor of the computer 48. This direct response to the user's computer is represented by route 50.

If the static document is not contained within the cache contents 46 of the cache access point 44 then the request must be connected directly to the web site 52 where the static document 54 is retrieved. The static document 54 is then routed through the cache access point 44 such that the static document 54 becomes cached in the cached contents 46 and then may be routed to the user's computer 48 through route 50.

If the inquiry of the cache access point 44 as to whether the web page of request 42 is static is answered in the negative, a second inquiry is undertaken. The second inquiry is whether the web page of request 42 contains any static elements as component parts of the web page.

If the web page of request 42 contains no static components or elements then the request is processed in a

manner similar to the process of the prior art system 10 of Figure 1 for responding to a web page request for a web page that is not completely static. That is, system 40 processes request 42 for a web page that consists of solely dynamic data by connecting the request 42 directly to the web site 52 with the resulting web page being routed directly to the user's computer 48 through route 56.

A third option available through the system 40 of the present invention is where a web page consists of both dynamic data as well as static elements or components. A web page of request 42 that includes these two types of information (i.e., dynamic and static) is parsed into its distinct components. That is, the dynamic data in the form of an XML (extendable markup language) document is processed as dynamic data that must be retrieved directly from the web site 42. The static elements, often referred to XSL (extendable scripting language) or CSS (cascading style sheets) data, however, may be contained in the cache contents 46 of the cache access point 44. Thus, a third inquiry is made, that being whether the XSL/CSS data is contained in the cache contents 46 of the cache access point 44. If the XSL/CSS data is not contained in the cache contents 46, then the request for the XSL/CSS data must be made directly to the web site 52 and then rerouted to the cache access point 44 such that the XSL/CSS data can be cached in the cache contents 46. At this point,

the XSL/CSS data of the cache contents 46 can be routed to the XML Parser 58. While the XSL/CSS data request is being processed, the XML information is simultaneously being retrieved from the web site 52 and routed to the XML Parser 58. At this time, the XML data and the XSL/CSS data may be combined by the XML Parser 58 and routed to the user's computer 48 through route 60.

The result of the use of an XML Parser 58 at the cache access point 44 is to increase the speed at which the internet can respond to a user's request 42. This is particularly evident for requests 42 for web pages that include both dynamic and static elements or components because the XML Parser 58 located at the cache access point 44 reduces the amount of data that must be directly obtained from the web site 52. That is, those static elements or components of the web page of request 42 may not need to be retrieved from the web site 52 leaving only the dynamic elements for retrieval. This system also further reduces the activity at the web site 52 thus enabling the web site 52 to process more requests. Finally, the system reduces the overall traffic of the internet and its various connections thus increasing the overall speed of all requests whether fully dynamic, fully static, or a combination of dynamic and static.

More generally, although the present invention has been described in connection with the preferred form of practicing it, those of ordinary skill in the art will

understand that many modifications can be made thereto within the scope of the claims that follow. Accordingly, it is not intended that the scope of the invention in any way be limited by the above description, but instead be determined entirely by reference to the claims that follow.

What is claimed:

1. A method for universally caching and distributing information from dynamically-generated websites comprising:
  - (a) transmitting a request for a webpage from a website;
  - (b) determining whether the requested webpage includes both static information and dynamic information;
  - (c) parsing the webpage that includes both static and dynamic information into its static component and its dynamic component;
  - (d) retrieving the static component of the requested webpage from a cache access point;
  - (e) retrieving the dynamic component directly from the website;
  - (f) combining the cached static component with the retrieved dynamic component; and
  - (g) distributing the combined cached static component and the retrieved dynamic component to the location from which the request was first transmitted.
2. A system for universally caching and distributing information from dynamically-generated websites comprising:
  - (a) an input device capable of receiving a request for a webpage from a website;
  - (b) a cache access point capable of processing the

request for a webpage by (i) determining whether the requested webpage includes both static information and dynamic information, (ii) parsing the webpage that includes both static and dynamic information into its static component and its dynamic component, (iii) retrieving the static component of the requested webpage from a cache access point, (iv) retrieving the dynamic component directly from the website, (v) combining the cached static component with the retrieved dynamic component, and (vi) distributing the combined cached static component and the retrieved dynamic component to the location from which the request was first transmitted; and

(c) a display for displaying the distributed combined cached static component and the retrieved dynamic component.

## Prior Art

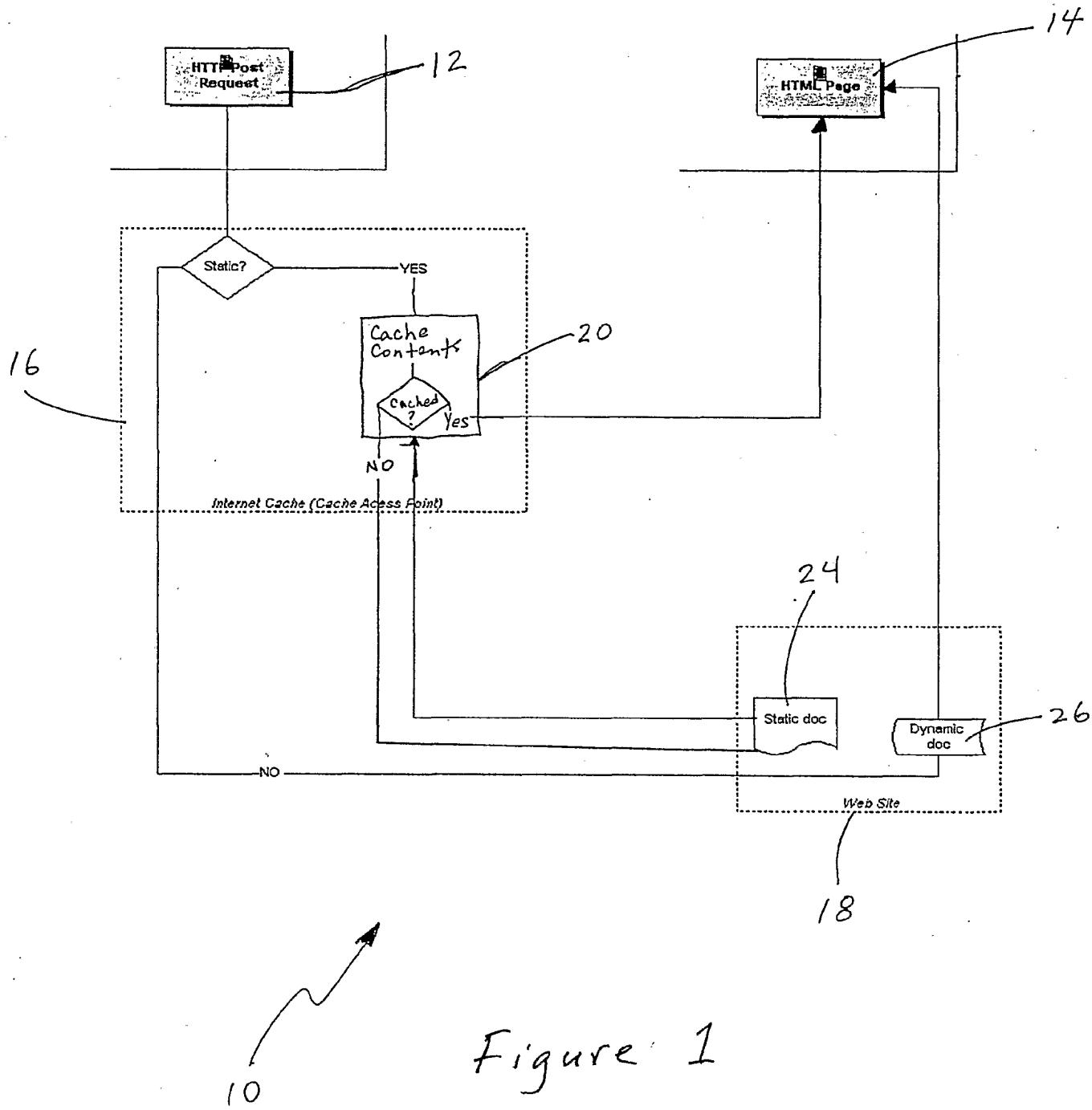
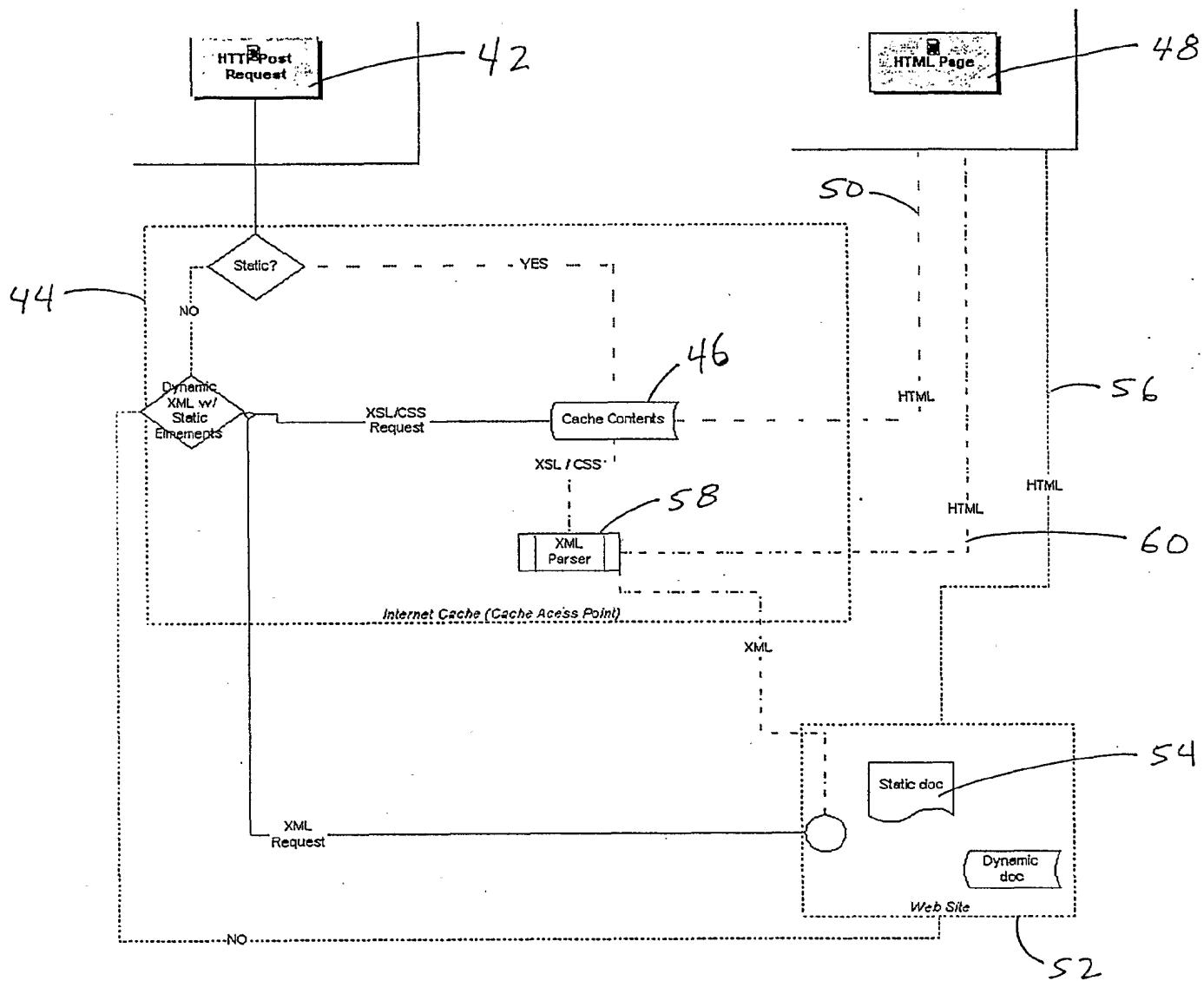


Figure 1



40  
↗

Figure 2

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US02/00268

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 19/00

US CL : 709/208

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 709/208; 711/100,117,118,126

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
**NONE**

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST

search terms: (cach\$ near5 (webpage\$ or (web adj1 page\$))) and (cach\$ near3 static)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6,021,426 A (DOUGLIS et al.) 01 February 2000 Abstract, figures 1-5, col. 2 (lines 30-52), col. 3 (lines 9-13, 30-44), and col. 6 (lines 44-56).	1 and 2
A	US 6,061,715 A (HAWES) 09 May 2000 see Abstract, figures 1-6, and col. 2 (line 15-et seq.).	1 and 2

Further documents are listed in the continuation of Box C.

See patent family annex.

"	Special categories of cited documents:	"T"	Later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

10 APRIL 2002

Date of mailing of the international search report

07 MAY 2002

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231  
Facsimile No. (703) 305-9290

Authorized officer *Robert Harrell*  
ROBERT B. HARRELL  
Telephone No. (703) 305-9692